DOCUMENT RESUME

ED 109 146 TM 004 575

AUTHOR Vitaliano, Peter Paul
TITLE Small Sample Comparisons of the Cochran Q and the Minimum X sub one squared Statistics.
PUB DATE [Apr 75]
NOTE 26p.; Paper presented at the Annual Meeting of the American Educational Research Association (Washington, D.C., March 30-April 3, 1975)

EDRS PRICE MF-$0.76 HC-$1.95 PLUS POSTAGE
DESCRIPTORS *Comparative Analysis; Correlation; *Hypothesis Testing; Matched Groups; *Nonparametric Statistics; *Statistical Analysis; *Tests of Significance
IDENTIFIERS *Chi Square; Cochran Q; Correlated Proportions

ABSTRACT

The Cochran Q and the Minimum X sub one squared statistics are two ways to test a hypothesis of equivalent correlated proportions. This study investigated the small sample properties of Q and X sub one squared by Monte Carlo methods. The observed distributions were compared for their rates of covergence to the limiting theoretical X sub one squared distribution, and for the degree to which their error rates approximated the nominal error rates. These latter comparisons allowed for idiosyncrasies that exist in the Q test of the correlated proportion hypothesis. Results show that the X sub one squared statistics is more powerful than the Q statistic for testing hypotheses of equivalent correlated proportions. (Author)

ED109146

TM 004 575

SMALL SAMPLE COMPARISONS OF THE COCHRAN Q AND THE MINIMUM X_1^2 STATISTICS

Peter Paul Vitaliano

Syracuse University

Paper presented at the Annual Meeting of the American Education Research Association

Washington, D. C.

April 2, 1975

Session # 22.21

Introduction

The purpose of this study was to compare the relative merits, for small samples, of the Cochran Q statistic (Cochran, 1950) and the minimum X_1^2 statistic (Neyman, 1949). These two statistics can be meaningfully compared since they not only represent alternative ways for testing hypotheses about the equality of correlated proportions, but they are lso both asymptotically distributed as χ^2.

Hypotheses about the equality of correlated proportions are the result of situations in which categorical data are obtained from either matched samples, or from repeated observations on subjects from one sample. Since it is assumed, in Q, that observations are dichotomously scored, the two category (c=2) per response situation was considered in this study. Table 1 contains such a situation, where i=1,..., n matched groups (or subjects, for the repeated measures situation); j=1,..., r matched subjects within group i (or r responses per subject i); and the response x_{ij} = 1 or 0 for a success or failure.

....

Insert Table 1 about here

....

In Table 1, the hypothesis of equal correlated proportions is

$$H_o: E(T_1/n) = E(T_2/n) = \ldots = E(T_r/n), \tag{1}$$

where T_j is the number of successes in the j-th column. Thus the implication in (1) is that the probability of a success is equal for all r treatments.

Numerous examples of the experimental situation in Table 1 are cited in standard statistical texts (Hays, 1963; Siegel, 1956; Winer, 1971). One may have a need to test a hypothesis, as in (1), in a number of content areas: in psychometrics one may wish to test whether the items on a test differ in difficulty. In social psychology, only one item or question may be of interest; in this

situation one may wish to test whether or not a specific group changes its response to this question over time. In clinical psychology, rater reliability may be a concern; one might ask - is the proportion of patients rated positively identical for different therapists? Finally, in experimental or comparative psychology one may wish to test whether the proportion of positive responses, in a particular species, is constant for different drugs (levels) or different doses of the same drug.

Theoretical Development of the Q and X_1^2 Statistics

In developing a test for (1), Cochran assumed that the u_i total for any row i is fixed, and that the expected value of x_{ij}, or the probability of a success p_{ij}, is constant for all r columns in row i. Thus:

$$E(x_{ij}) = p_{ij} = u_i/r, \text{ and} \tag{2}$$

$$V(x_{ij}) = (u_i/r)(1 - u_i/r). \tag{3}$$

With u_i fixed, $E(u_i) = u_i$ and $V(u_i) = 0$ and, since $V(x_{ij})$ is constant for all r cells in row i,

$$V(u_i) = rV(x_{ij}) + \sum_{j \neq k}^{r}\sum^{r} Cov(x_{ij},x_{ik}) = 0 \tag{4}$$

Cochran further assumed that $Cov(x_{ij},x_{ik})$ is constant for all $j \neq k$, thus from (4),

$$Cov(x_{ij},x_{ik}) = \frac{-(u_i/r)(1 - u_i/r)}{r - 1} \tag{5}$$

The results in (2), (3) and (5), along with the assumption that the i rows are independent, allowed Cochran to obtain the:

$$E(T_j) = \sum_{i=1}^{n} E(x_{ij}) = \sum_{i}^{n} u_i/r = \sum_{j}^{r} T_j/r = \bar{T}, \tag{6}$$

$$V(T_j) = \sum_{i}^{n} V(x_{ij}) = \sum_{i}^{n} (u_i/r)(1 - u_i/r), \tag{7}$$

and

$$Cov(T_j,T_k) = \frac{-\sum_{i}^{n}(u_i/r)(1 - u_i/r)}{r - 1}, \quad j \neq k. \tag{8}$$

Given these results, Cochran defined Q as:

$$Q = \frac{\sum_j^r (T_j - \bar{T})^2}{\frac{\sum_i (u_i)(1 - u_i/r)}{r - 1}}, \tag{9}$$

where the denominator in (9) can be shown to be equivalent to (7) - (8).

If one assumes that n is large, the T_j totals may be expected to tend to multivariate normality with common variance (7) and common covariance (8). Given that (1) is true, Cochran cited Walsh (1947) as having proven that if n is large, the ratio in Q will be distributed as χ^2 with r - 1 degrees of freedom. However, if one reviews an assumption made in deriving Q (specifically in passing from (4) to (5)), it becomes obvious that Q does not test (1) alone. Instead, Q simultaneously tests (1) and the hypothesis of equal population covariances, namely:

$$\underline{H}_o\text{: all } r(r-1) \text{ covariances are equal,} \tag{10}$$

or that all values of (8) are equal.

It should be apparent that (1), together with (10) is a much more exacting hypothesis than (1) alone. This is at least one problem the researcher faces when Q is used to test (1). Another problem one faces when using Q can be observed in the denominator of (9). It should be obvious that rows where u_i is either equal to zero or to r do not contribute to the value of Q. Moreover, since Q is insensitive to these rows one can never estimate, a priori, what power the Q test will have since the effective sample size of Q will be less than n. That is, whenever the probability of obtaining a u_i = 0 or r is not equal to zero, sample size attrition will occur. It shall be shown below that X_1^2 is affected by the original sample size.

The minimum X_1^2 statistic is an alternative to the Q statistic for testing (1). Neyman (1949) wrote the foundation paper which defined and theoretically developed this statistic. Since then several authors, namely Berkson (1955), Bhapkar (1961, 1965) and Grizzle et al (1969) have reformulated the X_1^2 statistic so that it could be easily applied in a wide variety of situations.

Grizzle's approach is most appealing for a number of reasons: first, categorical data are expressed in terms of familiar linear models. Second, the parameters of these models are then estimated and hypotheses about them are tested using the well known method of weighted least squares. Finally, because the linear model approach can be applied in so many situations, Grizzle's method represents a unified approach, both conceptually and computationally, for applying the X_1^2 statistic to categorical data. Because of these advantages, Grizzle's approach for testing (1) is the one which is cited in this paper.

Grizzle begins his discussion by referring to a general categorical situation; this scheme is depicted in Table 2a, where the rows represent s multinomial populations and the columns represent c categories of a response. Table 2b contains definitions of the notation that is employed in the development of the X_1^2 statistic.

....

Insert Table 2a-b about here

....

Given this notation, Grizzle defines a function:

$f_m(\underline{\pi}) = \underset{1xcs}{a_m} \; \underset{csx1}{\underline{\pi}}$, where $f_m(\underline{\pi})$ is assumed to be any function of the elements of $\underline{\pi}$ that has partial derivatives up to the second order with respect to π_{ij} and $m = 1,2,\ldots,u \leq v = s(c-1)$. For all u functions, one can write the general linear model as:

$$\underset{ux1}{F(\underline{\pi})} = \underset{uxcs}{A} \; \underset{csx1}{\underline{\pi}} , \qquad (11)$$

where A is a matrix of desired weights, with rank equal to u, so that one obtains u linearly independent $f(\underline{\pi})$'s.

Given (11), S is then defined as the sample estimate of the covariance matrix of $F(\underline{p})$: $\underset{uxu}{S} = \underset{uxcs}{A} \; \underset{csxcs}{V(\underline{p})} \; \underset{csxu}{A'}$, with the rank of S equal to u. It should be mentioned that if any frequency (n_{ij}) is zero, S will be singular;

in this case Berkson (1955) has recommended that n_{ij} be replaced by $1/c$, so that p_{ij} is $1/(c\ n_{i.})$.

Grizzle assumes further that the $f_m(\underline{\pi})$'s in (11) can be described in terms of a matrix X and a vector of unknown parameters $\underline{\beta}$:

$$\underset{ux1}{F(\pi)} = \underset{uxcs}{A}\ \underset{csx1}{\underline{\pi}} = \underset{uxw}{X}\ \underset{wx1}{\underline{\beta}}\ , \qquad (12)$$

where X is a type of design matrix of rank $w \leq u$. Based on (12), one can write the sum of squared deviations, of observed versus expected linear functions, as:

$$(F(\underline{p}) - X\,\underline{\beta})'\, S^{-1}\, (F(p) - X\,\underline{\beta}). \qquad (13)$$

Given (13), Grizzle then cites Neyman (1949, Theorem 4) as having proven that if an expression as in (13) is minimized with respect to $\underline{\beta}$, the result:

$$\underline{b} = X'S^{-1}X)^{-1}X'S^{-1}F(\underline{p})$$

will contain B.A.N. estimates of the parameters in $\underline{\beta}$, where B.A.N. estimates are asymptotically normal, efficient and consistent. Furthermore, the minimum value of (13) that is obtained will, according to Bhapkar (1961, Theorem 3), be the minimum X_1^2 test statistic for the fit of the model in (12), namely:

$$\text{the SS(due to the } \underline{H}_o: F(\underline{\pi}) = X\,\underline{\beta}) = F(\underline{p})'S^{-1}F(\underline{p}) - \underline{b}'(X'S^{-1}X)\underline{b}\ . \qquad (14)$$

Finally, since (14) is the minimum X_1^2 statistic it will, according to Neyman (1949, lemma 12), be asymptotically distributed as X^2 with u-w degrees of freedom, if (12) is true.

Once a model is defined and tested for adequate fit to the data (as in (14)), a test of a general linear hypothesis:

$$\underline{H}_o: \underset{dxw}{C}\ \underset{wx1}{\underline{\beta}}\ \underset{dx1}{0} \qquad (15)$$

can be obtained by the same weighted least squares method. In (15), C(dxw) is of full row rank $d \leq w$. Thus (15) represents restrictions on the original parameters in (12). Using the same rationale as in (13) and (14), one obtains a

minimum X_1^2, test statistic for (15) as:

$$SS(C\underline{\beta} = \underline{0}) = \underline{b}'C'(C(X'S^{-1}X))^{-1}C')^{-1}C\underline{b},$$

with d degrees of freedom, if (15) is true.

In many cases there is only one population, as in Table 1, and the objective of the analysis is to study the relationships among ways of classification of the sample units. Many situations of this kind can be described by the model:

$$\underset{u\times 1}{F(\underline{\pi})} = \underline{0} \ . \qquad (16)$$

This fits into Grizzle's general model (12) by setting $X = 0$, the null matrix. The test statistic for (16) is then:

$$SS(F(\underline{\pi}) = \underline{0}) = F(p)'S^{-1}F(\underline{p}), \qquad (17)$$

which is asymptotically χ^2 with u degrees of freedom, if (16) is true.

Since the situation in Table 1 is a one population problem, the expression in (17) can be used as the minimum X_1^2 test for the hypothesis in (1). This can be presented through the use of an example. Table 3 contains a reproduction of the repeated measures data presented in Grizzle's Table 4 (1969). Forty-six subjects were each given drugs A, B, and C. Some had a favorable response to one, some to two and some to all three drugs.

....

Insert Table 3 about here

....

Before the X_1^2 method for handling this situation is discussed one should note the differences between the situations in Tables 2a and 3: the former Table contains s populations, 1 response per subject and c categories per response, while the latter table contains 1 population, r responses per subject and 2 categories per response. However, although this latter situation has two categories it is still multinomial if one regards the possible response patterns for any one subject as the experiment's mutually exclusive categories. That is, for each of the three responses in Table 3 there are two possible categories,

therefore, there are a total of $c^r(=2^3=8)$ possible response patterns. Since any one subject can have one and only one response pattern his response vector contains eight elements, one of which is equal to one and seven of which are equal to zero.

Given this configuration the hypothesis to be tested in Table 3 is that the marginal proportions are equal:

$$\underline{H}_0 : E(T_1/N) = E(T_2/N) = E(T_3/N) \text{ or} \qquad (18)$$
$$\underline{H}_0 : E(T_1) = E(T_2) = E(T_3),$$

where (18) is identical to (1), for three responses.

This hypothesis may be written so that it is ammenable to Grizzle's approach. Given Table 3, one can see that (18) implies the hypothesis:

$$\underline{H}_0: \pi_1+\pi_2+\pi_3+\pi_5 = \pi_1+\pi_2+\pi_4+\pi_6 = \pi_1+\pi_3+\pi_4+\pi_7 .$$

Yet this hypothesis can be rewritten as:

$$\underline{H}_0: \begin{array}{r} \pi_1 + \pi_2 + \pi_3 + \pi_5 \\ -\pi_1 - \pi_3 - \pi_4 - \pi_7 \\ \hline \pi_2 - \pi_4 + \pi_5 - \pi_7 \end{array} = 0, \quad \begin{array}{r} \pi_1 + \pi_2 + \pi_4 + \pi_6 \\ -\pi_1 - \pi_3 - \pi_4 - \pi_7 \\ \hline \pi_2 - \pi_3 + \pi_6 - \pi_7 \end{array} = 0, \text{ or}$$

$$\underline{H}_0: \begin{array}{l} \pi_2 - \pi_4 + \pi_5 - \pi_7 = 0 \\ \pi_2 - \pi_3 + \pi_6 - \pi_7 = 0 \end{array} \qquad (19)$$

This hypothesis can be readily adapted to that in (16) by choosing A such that

$$A = \begin{bmatrix} 0 & 1 & 0 & -1 & 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 & 0 & 1 & -1 & 0 \end{bmatrix},$$

and (19) then becomes

$$\underline{H}_0: \quad F(\underline{\pi}) = A\,\underline{\pi} = \underline{0} . \qquad (20)$$

From (20), the estimated covariance matrix of $F(\underline{p})$ is $A\,V(\underline{p})A'$, and the test of the fit of the model in (20) is given by:

$$X_1^2 = \underline{p}'A'(AV(\underline{p})A')^{-1}A\underline{p} , \qquad (21)$$

where if (20) holds, (21) is asymptotically distributed as χ^2 with two degrees

of freedom. Table 4 provides the computations which are necessary for obtaining Q (in (9)) and X_1^2 (in (21)) for those data in Table 3.

....
Insert Table 4 about here
....

Given the theoretical development of the Q and X_1^2 statistics, it should be clear that while Q tests (1) and (10) simultaneously, X_1^2 only tests (1). Now that the foundations of these statistics have been presented, the methodology employed in this study shall be discussed.

Methodology

The properties of the small sample distributions of Q and X_1^2 were investigated (for $r = 2$) through the use of enumeration methods. Once a specific parent population was defined these two statistics were calculated on all possible samples from this population so that exact distributions of Q and X_1^2 could be formed.

Given the hypothesis that Q and X_1^2 test, there are four possible combinations of (1) true or false and (10) true or false. These combinations define four distinct population types; they are presented in Table 5. One should notice that the number of responses is limited in certain population types: if $r = 2$, then (10) can not be false and so population types B and C only contain $r > 2$ responses. Also, if $r > 2$, (1) can not be simultaneously false while (10) is true; therefore, type D populations only contain two responses.

....
Insert Table 5 about here
....

The behavior of Q and X_1^2 in three response populations is currently being investigated. A decision has not yet been reached as to whether or not enumeration will be possible in the three response case. Patil (1975) has provided a method for enumerating Q for any r; however, it is not yet clear

whether X_1^2 can be as easily enumerated in the r=3 situation. If enumeration is found to be intractable, the approximate distribution of X_1^2 will be simulated through the use of Monte Carlo methods.

It should be mentioned that Tate and Brown (1964, 1970) also studied the effect of differing sample sizes on the distribution of Q. They enumerated the exact distribution of Q for selected values of r and n . However, since these authors studied Q alone, and did not relate it to any other statistics, they did not find it necessary to generate distributions from populations in which sample size deletion occurs (where the probabilities of trivial patterns are non-zero). Furthermore, in studying Q in isolation, Tate and Brown only had to enumerate distributions of Q in type A populations. In the present study the distributions of Q and X_1^2 were generated from type A and D populations. Moreover, these populations were predetermined so that the probability in the trivial patterns varied. In this way, the effect that sample size attrition has on the discrepancies in the exact powers of Q and X_1^2, for any sample size, could be studied.

Results and Discussion

Before any results are reported, an important relationship between Q and X_1^2 (for r = 2) should be mentioned; once this relationship was established there was no doubt that the small sample power of X_1^2 would be greater than the small sample power of Q (at least for r = 2).

The X_1^2 statistic will now be shown to be monotonically related to Q. In the two response case there are $M = 2^r = 2^2 = 4$ possible patterns. If these patterns are denoted as: (11), (10), (01), (00) and their respective frequencies as: n_1, n_2, n_3, n_4, then it can be shown that for all nontrivial cases ($n_2 \neq n_3$), X_1^2 will always be greater than Q.

Given the above pattern frequencies, it is well known that Q (here identical to the McNemar statistic; McNemar, 1949) can be expressed as:

$$Q = \frac{(n_2 - n_3)^2}{n_2 + n_3} \quad (22)$$

Furthermore, if one manipulates the expression in (21), it can be shown that for r=2, (21) can be expressed as:

$$X_1^2 = \frac{(n_2 - n_3)^2}{n_2+n_3 - \frac{(n_2-n_3)^2}{N}} \quad (23)$$

Given (22) and (23) one can show that:

$$\frac{1}{X_1^2} = \frac{1}{Q} - \frac{1}{N}, \text{ or} \quad (24)$$

$$X_1^2 = Q\left(\frac{N}{N-Q}\right) \quad (25)$$

From (24) one can observe that (for all cases where $n_2 \neq n_3$) X_1^2 will always reject (1) more often than Q. The results below are consistent with this finding.

The exact α levels of Q and X_1^2

Table 6 contains the $P(Q \geq \chi^2_{.01,1})$, $P(X_1^2 \geq \chi^2_{.01,1})$, $P(Q \geq \chi^2_{.05,1})$ and $P(X_1^2 \geq \chi^2_{.05,1})$. These probabilities were obtained from exact sampling distributions based on three type A populations. A total of nine sampling distributions were generated from these populations; these distributions are ordered according to the actual sample size used in the calculation of Q (from the smallest to the largest N)

Insert Table 6 about here

In all eighteen comparisons of Q and X_1^2, the exact probability of Q is closer to the asymptotic nominal probability (based on the χ^2 distribution) than

is the exact probability of X_1^2. Q is conservative in fifteen of the eighteen comparisons (in all nine at $\alpha = .01$), while X_1^2 is liberal in all eighteen comparisons. As N increases, the upper tails of both the Q and X_1^2 distributions converge to the upper tail of the χ^2 distribution, although Q seems to converge more rapidly. Moreover, as expected from (24), as N increases the exact distributions of Q and X_1^2 approach one another.

Based on the results in Table 6 one can conclude that for: $r = 2$, $N = 40$ and (1) and (10) true, the distribution of Q comes closer to the χ^2 distribution than does the distribution of X_1^2; thus in these situations, Q can be said to provide a better test than X_1^2. In general the discrepancies of these two statistics are not as great for $\alpha = .05$ as for $\alpha = .01$.

The exact power of Q and X_1^2.

Table 7 contains thirty two exact distributions which were enumerated from twenty-five different type D populations. The distributions are ordered according to the magnitude (from smallest to largest) of the noncentrality parameter (λ) in their respective parent populations. The estimated asymptotic power in any population is also given; this was obtained by referring the appropriate λ, α and degrees of freedom to tables of the noncentral χ^2 distribution (Owen, 1964). The estimated asymptotic power in a parent population serves as a reference point for comparing the exact powers of Q and X_1^2; that is, if the asymptotic power in a parent population were very high (say .95) it would be difficult to detect differences in the exact powers of Q and X_1^2, since these two statistics would both be quite powerful. It was arbitrary as to whether the approximate asymptotic power of Q or X_1^2 would be used as the reference point. It was decided that the asymptotic power of X_1^2 be used and so the λ for a particular population was obtained by substituting, into (23), the values of $N\pi_2$ and $N\pi_3$ in that population.

....
Insert Table 7 about here
....

As expected from (24), χ_1^2 is consistently more powerful than Q; in all sixty-four comparisons, $P(\chi_1^2 \geq \chi_\alpha^2)$ is greater than $P(Q \geq \chi_\alpha^2)$. One might point out a fallacy in this type of comparison. That is, one might argue that while Q is less powerful than χ_1^2, it is also a much more conservative test. In this sense, it would appear to be unfair to compare the exact powers of Q and χ_1^2 by comparing $P(Q \geq \chi_\alpha^2)$ to $P(\chi_1^2 \geq \chi_\alpha^2)$. If, on the other hand, one were to find the values of Q=K and χ_1^2 = L such that $\alpha = P(Q \geq K) = P(\chi_1^2 \geq L)$, then by definition, the exact powers of Q and χ_1^2 would also be identical. The problem with this argument is that the researcher does not know the exact distribution of either Q or χ_1^2 (he cannot obtain K or L), and in using the χ^2 approximation, he assumes that $P(Q \geq \chi_\alpha^2) = P(\chi_1^2 \geq \chi_\alpha^2) = \alpha$. Given this assumed constant value of χ_α^2, (1) will be rejected more often with χ_1^2 than with Q, and in this sense it is legitimate to say that χ_1^2 is more powerful than Q. The logical question is then - in which situations would one statistic be preferred over the other? The trivial response is that Q is to be preferred when type I errors are of major concern, while χ_1^2 is to be preferred when type II errors are of major concern. However, a more specific set of recommendations can be offered if one is willing to make a concession.

Suppose a researcher has a sample size of N = 20 or N = 30; if he were to use the χ_1^2 statistic to test a hypothesis, the exact type I error rate, at asymptotic α = .01, would only be partially controlled if N were 20 (in Table 6 the average error rate for N=20 is .045); however, if n were 30, the exact error rate would be controlled adequately (the average at N=30 is .0185). Given this scheme one might ask, if these exact error rates were tolerated in χ_1^2, under what conditions would the exact powers of χ_1^2 and Q be most discrepant? If this question could be answered, then some valuable recommendations could be offered to researchers who were willing to tolerate the above type I error rates. The results in Table 8 provide a partial answer to this question.

....

Insert Table 8 about here

....

Table 8 represents a re-examination of the results obtained for the twenty exact distributions (in Table 7) which were generated from samples of $N = 20$ or $N = 30$. The rows and columns in this table respectively represent the asymptotic powers and percent deletions in the parent populations of each of these twenty distributions. The cells in the body of the table contain the following information: the label of the distribution, as in Table 7; in parenthesis, the ratio of the difference in the exact powers of X_1^2 and Q, divided by the exact power of X_1^2 (hereafter this ratio will be referred to as RDP); and finally, the sample size used to obtain the distribution. It should be mentioned that RDP is less affected by the asymptotic power in a parent population than is a mere difference in the exact powers and, in this sense, relative differences are more informative than are absolute differences.

If one observes the RDP's in Table 8 it should be evident that while X_1^2 is always recommended over Q (in terms of having greater power), the use of X_1^2 is almost a necessity in certain situations. As RDP approaches zero, the exact powers of X_1^2 and Q approach one another. Given this, it should be clear that under certain circumstances the use of X_1^2 is highly recommended over the use of Q: very large RDP's occur when asymptotic power is low and the percent deletions is large (populations 1-3, 7 and 9); however, if asymptotic power is great, even large deletions do not appreciably affect the RDP (populations 29 and 33).

Given a range of low asymptotic power (all populations whose noncentrality parameters are less than 5), it is clear that as the percent deletion increases, RDP increases dramatically. In fact, populations 8 and 12 have more than three

times the asymptotic power of populations 1 and 2, yet in spite of these differences, the larger percentage deletion in 1 and 2 produces RDP's which are much larger than in 8 and 12. One very telling result is in the difference in the RDP's of 8 and 9. These populations have identical asymptotic power, yet 9 has more than twice the number of deletions as 8; simultaneously, the RDP in 9 is twice as large as in 8.

Based on the results in Table 8, one can make thi recommendation: if a researcher is willing to tolerate the exact α levels of X_1^2 when N = 20 or 30 (the latter level is certainly tolerable) then X_1^2 should always be used in place of Q ; however, if the researcher obtains sample frequencies where the non-trivial frequencies are simultaneously close in value and small, say $|n_2 - n_3| < 3$ and the percent deletions is greater than .8, then the exact power of X_1^2 is so much greater than the exact power of Q that one should certainly use the X_1^2 statistic.

REFERENCES

Berkson, J. Maximum likelihood and minimum χ^2 estimates of the logistic function, Journal of the American Statistical Association, 1955, 50, 130-162.

Bhapkar, V. P. Some tests for categorical data. Annals of Mathematical Statistics, 1961, 32, 72-83.

Bhapkar, V. P. Categorical data analogs of some multivariate tests. S. N. Roy Memorial Volume, 1965, University of North Carolina Press, Chapel Hill, NC

Cochran, W. G. The comparison of percentages in matched samples. Biometrika, 1950, 37, 256-66.

Grizzle, J. L., Starmer, C. F., and Koch, G. G. Analysis of categorical data by linear models. Biometrics, 1969, 25, 489-504.

Hays, W. L. Statistics for Psychologists. New York: Holt, Rinehart and Winston, Inc., 1963.

Mc Nemar, Q. Psychological Statistics. New York: John Wiley, 1949.

Neyman, J. Contribution to the theory of the χ^2 test. Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, 1949, 239-273.

Owen, D. B. Handbook of statistical tables. Reading, Mass.: Addison-Wesley, 1962.

Patil, K. Cochran's Q test: Exact distribution. Journal of the American Statistical Association, 1975, in press.

Siegel, S. Nonparametric statistics for the behavioral sciences. New York; Mc Graw-Hill, 1956.

Tate, M. W., and Brown, S. M. Tables for comparing related sample percentages and for the median test. Philadelphia: Graduate School of Education, University of Pennsylvania, 1964.

Tate, M. W., and Brown, S. M. Note on the Cochran Q test, Journal of the American Statistical Association, 1970, 65, 155-160.

Walsh, J. E. Concerning the effect of intraclass correlation on certain significance tests. Annals of Mathematical Statistics, 1947, 18, 88-

Winer, B. J. Statistical principles in experimental design. New York: McGraw-Hill, 1971.

Table 1 : A Matched Group (or Repeated Measures) Categorical Situation

		MATCHED SUBJECTS (OR REPEATED MEASURES FOR ONE SUBJECT)				
MATCHED GROUPS (OR SUBJECTS)		$j = 1$	2	. . . . . .	r	$\sum_j^r x_{ij} = u_i$
	$i = 1$	$x_{ij} = x_{11}$	x_{12}	. . . . .	x_{1r}	$u_i = u_1$
	2	x_{21}	x_{22}	. . . . .	x_{2r}	u_2
	. .	. .	. .	. . . .	. .	. .
	n	x_{n1}	x_{n2}	. . . . .	x_{nr}	u_n
	$\sum_i^n x_{ij}$	$= T_j = T_1$	T_2	. . . . .	T_r	$\sum_j T_j = \sum_i^n u_i$

Table 2a : General Categorical Situation for s Multinomial Populations

POPULATIONS	CATEGORIES OF RESPONSE				
	$j = 1$	2	...	c	TOTAL
$i = 1$	$\pi_{ij} = \pi_{11}$	π_{12}	...	π_{1c}	$\sum_{j}^{c} \pi_{ij} = 1$
2	π_{21}	π_{22}	..	π_{2c}	1
..	..	..	...	..	..
s	π_{s1}	π_{s2}	...	π_{sc}	1

Table 2b : The Notation used in the Development of the X_1^2 Statistic

$$\underset{1\times c}{\underline{\pi}_i'} = (\pi_{i1}, \pi_{i2}, \ldots, \pi_{ic})$$

$$\underset{1\times cs}{\underline{\pi}} = (\underline{\pi}_1', \underline{\pi}_2', \ldots, \underline{\pi}_s')$$

$$p_{ij} = \frac{n_{ij}}{n_{i.}} \; ; \; \underset{1\times c}{\underline{p}_i'} = (p_{i1}, p_{i2}, \ldots, p_{ic}) \; ; \; n_{i.} = \text{i-th sample size.}$$

$$\text{Variance } (\underline{p}_i) = \underset{c\times c}{V(\underline{\pi}_i)} = \frac{1}{n_{i.}} \begin{bmatrix} \pi_{i1}(1-\pi_{i1}) & -\pi_{i1}\pi_{i2} & \cdots & -\pi_{i1}\pi_{ic} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ -\pi_{ic}\pi_{i1} & -\pi_{ic}\pi_{i2} & \cdots & \pi_{ic}(1-\pi_{ic}) \end{bmatrix}$$

$$\underset{c\times c}{V(\underline{p}_i)} = V(\underline{\pi}_i), \text{ with } \pi_{ij} \text{ replaced by } p_{ij}.$$

$$\underset{cs\times cs}{V(\underline{p})} = \text{block diagonal matrix of } V(\underline{p}_i)$$

Table 3 : An Example of data obtained from a Repeated Measures Experiment (from Grizzle, 1969).

		TABULATION OF RESPONSES TO DRUGS A, B, AND C			FREQUENCY OF SUBJECTS	EXPECTED PROBABILITY
		PATTERN OF RESPONSES			n_j	p_j
		A	B	C		
Eight Possible Response Patterns	1	1	1	1	$n_1 = 6$	p_1
	2	1	1	0	$n_2 = 16$	p_2
	3	1	0	1	$n_3 = 2$	p_3
	4	0	1	1	$n_4 = 2$	p_4
	5	1	0	0	$n_5 = 4$	p_5
	6	0	1	0	$n_6 = 4$	p_6
	7	0	0	1	$n_7 = 6$	p_7
	8	0	0	0	$n_8 = 6$	p_8
		Number Favorable			$\sum_{j}^{8} n_j = N = 46$	$\sum_{j}^{8} p_j = 1$
		$n_1 + n_2 + n_3 + n_5 = T_1 = 28$	$n_1 + n_2 + n_4 + n_6 = T_2 = 28$	$n_1 + n_3 + n_4 + n_7 = T_3 = 16$		

Table 4 : The Computations which are necessary for obtaining the values of the X_1^2 and the Q statistics for those data in Table 3.

The Computation of X_1^2 :

$$X_1^2 = \underline{p}'A' \, (A \, V(\underline{p}) \, A')^{-1} \, A\underline{p},$$

where $\underline{p}'A' = (.26, .26)$, and $(AV(\underline{p})A')^{-1} = \frac{46}{1241} \begin{bmatrix} .5406 & -.4101 \\ -.4101 & .5406 \end{bmatrix} 10^4$,

then $X_1^2 = 6.58$

The Computation of Q :

$$Q = \frac{(r-1) \, \Sigma(T_j - \bar{T})^2}{\Sigma u_i - \frac{\Sigma u_i^2}{r}} = \frac{2 \, [(28 - 24)^2 + (28 - 24)^2 + (16 - 24)^2]}{72 - \frac{148}{3}}$$

$$Q = \frac{2 \, (96)}{22.67} = \frac{192}{22.67} = 8.47$$

Table 5 : The Four Possible Population Types from which Distributions of Q and X_1^2 can be generated.

(10) H_o : all $\mathrm{cov}(T_j, T_k)$ are equal, $j \neq k$	(1) H_o : $E(T_1) = E(T_2) = \ldots = E(T_r)$ TRUE	(1) H_o : $E(T_1) = E(T_2) = \ldots = E(T_r)$ FALSE
TRUE	**Type A Populations** H_o is true for Q and X_1^2: for r = 2 and r = 3 1. Sampling distributions of Q and X_1^2 were both compared to the χ^2 distribution and to each other. Specific interest was in the exact versus the asymptotic α level of each test and the effect of sample size on the adequacy of the χ^2 approximation to the upper portions of the Q and X_1^2 distributions. 2. The effect of sample size attrition was studied in Q along with the effects of sample deletion on differences in the exact type I error rates of Q and X_1^2.	**Type D Populations** H_o is false for Q and X_1^2: for r = 2 1. The exact powers of Q and X_1^2 were studied in relation to asymptotic power, sample size and sample size deletion. 2. Specific interest was in the degree to which discrepancies in the exact power of Q and X_1^2 are affected by sample size deletion, sample size and asymptotic power.
FALSE	**Type C Populations** H_o is true for X_1^2 and false for Q.. r = 3 The following phenomena are currently under investigation: 1. The effect, on the distribution of Q, of the departure of (10) from equal covariances. 2. The effect of sample size and sample size attrition on the exact power of Q. 3. The exact α levels in the distributions of the X_1^2 statistic.	**Type B Populations** H_o is false for Q and X_1^2. r = 3 The same comparisons, as in Type D populations, are currently under investigation.

Table 6 : The Exact α Levels for Nine Type A Distributions (where the hypotheses in (1) and (10) are simultaneously true)

EXACT DISTRIBUTION	TYPE A POPULATION*	SAMPLE SIZE USED IN Q	ORIGINAL SAMPLE SIZE	PERCENT DELETION	ASYMPTOTIC α = .01		ASYMPTOTIC α = .05	
					$P(Q \geq 6.64)$**	$P(X_1^2 \geq 6.64)$	$P(Q \geq 3.84)$	$P(X_1^2 \geq 3.84)$
1	3	6	10	40	.0045	.0478	.0479	.0882
2	2	8	10	20	.0067	.0423	.0423	.1116
3	2	12	15	20	.0071	.0241	.0409	.0862
4	3	12	20	40	.0072	.0238	.0426	.0724
5	2	16	20	20	.0081	.0228	.0502	.0777
6	1	18	20	10	.0079	.0246	.0489	.0704
7	3	18	30	40	.0080	.0183	.0515	.0647
8	2	24	30	20	.0091	.0188	.0541	.0653
9	1	36	40	10	.0096	.0169	.0489	.0576

* 1, 2 and 3 respectively represent these populations:

1. $\pi_1 = .05 \quad \pi_2 = .45 \quad \pi_3 = .45 \quad \pi_4 = .05$

2. $\pi_1 = .2 \quad \pi_2 = .4 \quad \pi_3 = .4 \quad \pi_4 = 0$

3. $\pi_1 = .4 \quad \pi_2 = .3 \quad \pi_3 = .3 \quad \pi_4 = 0$

** $6.64 = \chi^2_{.01,1}$; $3.84 = \chi^2_{.05,1}$

Table 7 : Exact Distributions from Type D Populations ((1) false and (10) true)

	PATTERN PROBABILITIES 11 π_1	10 π_2	01 π_3	00 π_4	Non-Central Parameter λ	Approx. Asymot. Power	N used in Q	Actual N	% deletion	$\alpha = .01$ $P(Q \geq \chi^2_{.01,1})$	$P(\chi^2_1 \geq \chi^2_{.01,1})$	$\alpha = .05$ $P(Q \geq \chi^2_{.05,1})$	$P(\chi^2_1 \geq \chi^2_{.05,1})$
1	.7	.2	.1	0	.69	< .1	6	20	70	.0176	.0700	.1174	.1584
2	.8	.133	.067	0	.86	< .1	6	30	80	.0172	.0441	.1156	.1552
3	.933	.067	0	0	2.03	>.1	2	30	93.3	.0028	.0121	.1294	.1296
4	.2	.6	.2	0	2.5	>.1	8	10	20	.0833	.2605	.2605	.443
5	.15	.5	.2	.15	2.95	<.2	14	20	30	.1391	.2540	.3528	.4417
6	.6	.333	.067	0	3.24	>.2	6	15	60	.0959	.2688	.3758	.4521
7	.9	.1	0	0	3.33	>.2	3	30	90	.0252	.0715	.3435	.3447
8	.3	.45	.15	.10	3.53	>.2	12	20	40	.1633	.3032	.3947	.4849
9	.85	.15	0	0	3.53	>.2	3	20	85	.0216	.1674	.3463	.3471
10	.133	.6	.2	.067	3.75	>.2	12	15	20	.1621	.3049	.3887	.4871
11	.1	.7	.2	0	3.85	<.3	9	10	10	.1044	.3512	.3512	.5936
12	.467	.4	.133	0	4.61	>.3	16	30	46.7	.2606	.3600	.5346	.5819
13	.067	.533	.133	.267	4.74	>.3	10	15	33.3	.1996	.3727	.4738	.5748
14	.1	.6	.2	.1	5	<.4	16	20	20	.2612	.4059	.5402	.6211
15	.533	.367	.1	0	5.39	.4	14	30	53.3	.3097	.4102	.5882	.6482
16	.2	.6	.1	.1	5.56	>.4	7	10	30	.1950	.4880	.4880	.6373
17	.25	.55	.15	.05	5.93	>.4	14	20	30	.3105	.4774	.5903	.6823
18	.6	.333	.067	0	6.48	<.5	12	30	60	.3747	.4993	.6623	.7366
19	.3	.45	.15	.1	7.06	>.5	24	40	40	.4535	.5446	.7297	.7584
20	.2	.5	.1	.2	7.27	>.5	12	20	40	.3778	.5709	.6674	.7536
21	.8	.2	0	0	7.5	<.6	6	30	80	.3839	.5627	.8623	.8676
22	.667	.3	.033	0	8.13	>.6	10	30	66.7	.4952	.6572	.7863	.8540
23	.2	.7	.1	0	8.18	>.6	8	10	20	.2918	.6108	.6108	.7556
24	.2	.667	.133	0	8.28	>.6	12	15	20	.3796	.5780	.6685	.7541
25	.2	.6	.2	0	8.88	>.6	24	30	20	.4582	.5687	.7329	.7610
26	.2	.65	.15	0	9.09	<.7	16	20	20	.4808	.6431	.7617	.8235
27	.067	.533	.133	.267	9.47	<.7	20	30	33.3	.5571	.6822	.8206	.8457
28	.1	.6	.2	.1	10	>.7	32	40	20	.6211	.7059	.8425	.8604
29	.733	.267	0	0	10.9	>.7	8	30	73.3	.7125	.8418	.9677	.9717
30	.333	.6	.067	0	11.16	<.8	10	15	33.3	.4943	.7073	.7959	.8651
31	.2	.5	.1	.2	14.54	<.9	24	40	40	.8099	.8690	.9488	.9579
32	.2	.6	.1	.1	16.66	>.9	21	30	30	.8209	.8956	.9569	.9654

Table 8. The RDP's* (for $\alpha = .01$) of twenty exact sampling distributions based on sample sizes of twenty and thirty.

ASYMPTOTIC POWER	Noncentrality Parameter	THE PERCENTAGE OF DELETIONS IN THE Q STATISTIC						
		.20	30 - 33.3	40 - 46.6	53.3 - 60	66.6 - 73.3	80 - 85	90 - 93.3
< .1	.69					1(.75)20		
< .1	.86						2(.61)30	
< .2	2.03							3(.7686)30
< .2	2.95		5(.4528)30					
> .2	3.33							7(.6476)30
> .2	3.53			8(.442)20			9(.871)20	
< .3	4.61			12(.2761)30				
< .4	5.00	14(.3565)20						
.4	5.39			15(.245)30				
> .4	5.93		17(.3498)30					
< .5	6.48			18(.2495)30				
> .5	7.27		20(.3382)20					
< .6	7.50						21(.3177)30	
> .6	8.13					22(.2465)30		
> .6	8.88		25(.1943)30					
< .7	9.09	26(.2524)20						
< .7	9.47		27(.1833)30					
> .7	10.90						29(.1536)30	
> .9	16.66							33(.0834)30

* The RDP for each distribution is presented in parenthesis within each cell.